

## Patent Application US/07/659,408

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: CAPUT, DANIEL  
FERRARA, PASCUAL  
GUILLEMOT, JEAN-CLAUDE  
KAGHAD, MOURAD  
LEGOUX, RICHARD  
LOISON, GERARD  
LARBRE, ELIZABETH  
LUPKER, JOHANNES  
LEPLATOIS, PASCUAL  
SALOME, MARK

(ii) TITLE OF INVENTION: URATE OXIDASE ACTIVITY PROTEIN,  
RECOMBINANT GENE CODING THEREFOR, EXPRESSION VECTOR,  
MICRO-ORGANISMS AND TRANSFORMED CELLS

(iii) NUMBER OF SEQUENCES: 35

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Foley & Lardner  
(B) STREET: 1800 Diagonal Road, Suite 500  
(C) CITY: Alexandria  
(D) STATE: Virginia  
(E) COUNTRY: USA  
(F) ZIP: 22313-0299

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 07/659,408  
(B) FILING DATE: 25-APR-1991  
(C) CLASSIFICATION:

## (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: BENT, Stephen A.  
(B) REGISTRATION NUMBER: 29,768  
(C) REFERENCE/DOCKET NUMBER: 16781/276 BEDL

## (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (703)836-9300  
(B) TELEFAX: (703)683-4109  
(C) TELEX: 899149

## (2) INFORMATION FOR SEQ ID NO:1:

OK

## Patent Application US/07/659,408

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 301 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: NO

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: Aspergillus flavus

## (vii) IMMEDIATE SOURCE:

(B) CLONE: Urate oxidase

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Ser Ala Val Lys Ala Ala Arg Tyr Gly Lys Asp Asn Val Arg Val Tyr  
1 5 10 15

Lys Val His Lys Asp Glu Lys Thr Gly Val Gln Thr Val Tyr Glu Met  
20 25 30

Thr Val Cys Val Leu Leu Glu Gly Glu Ile Glu Thr Ser Tyr Thr Lys  
35 40 45

Ala Asp Asn Ser Val Ile Val Ala Thr Asp Ser Ile Lys Asn Thr Ile  
50 55 60

Tyr Ile Thr Ala Lys Gln Asn Pro Val Thr Pro Pro Glu Leu Phe Gly  
65 70 75 80

Ser Ile Leu Gly Thr His Phe Ile Glu Lys Tyr Asn His Ile His Ala  
85 90 95

Ala His Val Asn Ile Val Cys His Arg Trp Thr Arg Met Asp Ile Asp  
100 105 110

Gly Lys Pro His Pro His Ser Phe Ile Arg Asp Ser Glu Glu Lys Arg  
115 120 125

Asn Val Gln Val Asp Val Val Glu Gly Lys Gly Ile Asp Ile Lys Ser  
130 135 140

Ser Leu Ser Gly Leu Thr Val Leu Lys Ser Thr Asn Ser Gln Phe Trp  
145 150 155 160

Gly Phe Leu Arg Asp Glu Tyr Thr Thr Leu Lys Glu Thr Trp Asp Arg  
165 170 175

Ile Leu Ser Thr Asp Val Asp Ala Thr Trp Gln Trp Lys Asn Phe Ser

## Patent Application US/07/659,408

```

107          180          185          190
108
109 Gly Leu Gln Glu Val Arg Ser His Val Pro Lys Phe Asp Ala Thr Trp
110          195          200          205
111
112 Ala Thr Ala Arg Glu Val Thr Leu Lys Thr Phe Ala Glu Asp Asn Ser
113          210          215          220
114
115 Ala Ser Val Gln Ala Thr Met Tyr Lys Met Ala Glu Gln Ile Leu Ala
116          225          230          235          240
117
118 Arg Gln Gln Leu Ile Glu Thr Val Glu Tyr Ser Leu Pro Asn Lys His
119          245          250          255
120
121 Tyr Phe Glu Ile Asp Leu Ser Trp His Lys Gly Leu Gln Asn Thr Gly
122          260          265          270
123
124 Lys Asn Ala Glu Val Phe Ala Pro Gln Ser Asp Pro Asn Gly Leu Ile
125          275          280          285
126
127 Lys Cys Thr Val Gly Arg Ser Ser Leu Lys Ser Lys Leu
128          290          295          300
129 (2) INFORMATION FOR SEQ ID NO:2:
130
131 (i) SEQUENCE CHARACTERISTICS:
132 (A) LENGTH: 302 amino acids
133 (B) TYPE: amino acid
134 (C) STRANDEDNESS: single
135 (D) TOPOLOGY: linear
136
137 (ii) MOLECULE TYPE: protein
138
139 (iii) HYPOTHETICAL: NO
140
141 (vi) ORIGINAL SOURCE:
142 (A) ORGANISM: Aspergillus flavus
143
144 (vii) IMMEDIATE SOURCE:
145 (B) CLONE: Met-Urate oxidase
146
147
148 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
149
150 Met Ser Ala Val Lys Ala Ala Arg Tyr Gly Lys Asp Asn Val Arg Val
151 1          5          10          15
152
153 Tyr Lys Val His Lys Asp Glu Lys Thr Gly Val Gln Thr Val Tyr Glu
154          20          25          30
155
156 Met Thr Val Cys Val Leu Leu Glu Gly Glu Ile Glu Thr Ser Tyr Thr
157          35          40          45
158
159 Lys Ala Asp Asn Ser Val Ile Val Ala Thr Asp Ser Ile Lys Asn Thr

```

## Patent Application US/07/659,408

	50	55	60
160			
161			
162	Ile Tyr Ile Thr Ala Lys Gln Asn Pro Val Thr Pro Pro Glu Leu Phe		
163	65	70	75 80
164			
165	Gly Ser Ile Leu Gly Thr His Phe Ile Glu Lys Tyr Asn His Ile His		
166		85	90 95
167			
168	Ala Ala His Val Asn Ile Val Cys His Arg Trp Thr Arg Met Asp Ile		
169		100	105 110
170			
171	Asp Gly Lys Pro His Pro His Ser Phe Ile Arg Asp Ser Glu Glu Lys		
172		115	120 125
173			
174	Arg Asn Val Gln Val Asp Val Val Glu Gly Lys Gly Ile Asp Ile Lys		
175		130	135 140
176			
177	Ser Ser Leu Ser Gly Leu Thr Val Leu Lys Ser Thr Asn Ser Gln Phe		
178		145	150 155 160
179			
180	Trp Gly Phe Leu Arg Asp Glu Tyr Thr Thr Leu Lys Glu Thr Trp Asp		
181		165	170 175
182			
183	Arg Ile Leu Ser Thr Asp Val Asp Ala Thr Trp Gln Trp Lys Asn Phe		
184		180	185 190
185			
186	Ser Gly Leu Gln Glu Val Arg Ser His Val Pro Lys Phe Asp Ala Thr		
187		195	200 205
188			
189	Trp Ala Thr Ala Arg Glu Val Thr Leu Lys Thr Phe Ala Glu Asp Asn		
190		210	215 220
191			
192	Ser Ala Ser Val Gln Ala Thr Met Tyr Lys Met Ala Glu Gln Ile Leu		
193		225	230 235 240
194			
195	Ala Arg Gln Gln Leu Ile Glu Thr Val Glu Tyr Ser Leu Pro Asn Lys		
196		245	250 255
197			
198	His Tyr Phe Glu Ile Asp Leu Ser Trp His Lys Gly Leu Gln Asn Thr		
199		260	265 270
200			
201	Gly Lys Asn Ala Glu Val Phe Ala Pro Gln Ser Asp Pro Asn Gly Leu		
202		275	280 285
203			
204	Ile Lys Cys Thr Val Gly Arg Ser Ser Leu Lys Ser Lys Leu		
205		290	295 300
206			

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 906 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

## Patent Application US/07/659,408

213 (D) TOPOLOGY: linear  
214  
215 (ii) MOLECULE TYPE: DNA (genomic)  
216  
217 (vii) IMMEDIATE SOURCE:  
218 (B) CLONE: Preferred sequence for expression in  
219 prokaryotes  
220  
221  
222 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:  
223  
224 ATGTCTGCGG TAAAAGCAGC GCGCTACGGC AAGGACAATG TTCGCGTCTA CAAGGTTCAC 60  
225  
226 AAGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTGT GCTTCTGGAG 120  
227  
228 GGTGAGATTG AGACCTCTTA CACCAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC 180  
229  
230 ATTAAGAACA CCATTTACAT CACCGCCAAG CAGAACCCCG TTA CTCTCC CGAGCTGTTC 240  
231  
232 GGCTCCATCC TGGGCACACA CTTCA TTGAG AAGTACAACC ACATCCATGC CGCTCACGTC 300  
233  
234 AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC 360  
235  
236 TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC 420  
237  
238 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC 480  
239  
240 TGGGGCTTCC TCGGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC 540  
241  
242 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCA GTG GACTCCAGGA GGTCCGCTCG 600  
243  
244 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT 660  
245  
246 GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG 720  
247  
248 GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTTCGAA 780  
249  
250 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 840  
251  
252 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 900  
253  
254 AAATTG 906  
255  
256 (2) INFORMATION FOR SEQ ID NO:4:  
257 (i) SEQUENCE CHARACTERISTICS:  
258 (A) LENGTH: 906 base pairs  
259 (B) TYPE: nucleic acid  
260 (C) STRANDEDNESS: single  
261 (D) TOPOLOGY: linear  
262  
263 (ii) MOLECULE TYPE: DNA (genomic)  
264  
265

## Patent Application US/07/659,408

266 (vii) IMMEDIATE SOURCE:

267 (B) CLONE: Preferred sequence for expression in  
268 eukaryotes  
269270 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:  
271

272	ATGTCTGCTG TTAAGGCTGC TAGATACGGT AAGGACAACG TTAGAGTCTA CAAGGTTTAC	60
273		
274	AAGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTGT GCTTCTGGAG	120
275		
276	GGTGAGATTG AGACCTCTTA CACCAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC	180
277		
278	ATTAAGAACA CCATTTACAT CACCGCCAAG CAGAACCCCG TTAATCCTCC CGAGCTGTTT	240
279		
280	GGCTCCATCC TGGGCACACA CTTTATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC	300
281		
282	AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC	360
283		
284	TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC	420
285		
286	ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC	480
287		
288	TGGGGCTTCC TGCCTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC	540
289		
290	ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTTCACTG GACTCCAGGA GGTCCGCTCG	600
291		
292	CACGTGCCTA AGTTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT	660
293		
294	GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG	720
295		
296	GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTTGAA	780
297		
298	ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT	840
299		
300	CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT	900
301		
302	AAATTG	906
303		

304 (2) INFORMATION FOR SEQ ID NO:5:  
305

306 (i) SEQUENCE CHARACTERISTICS:

307 (A) LENGTH: 14 base pairs

308 (B) TYPE: nucleic acid

309 (C) STRANDEDNESS: single

310 (D) TOPOLOGY: linear  
311312 (ii) MOLECULE TYPE: DNA (genomic)  
313314 (iii) HYPOTHETICAL: NO  
315  
316

317 (vii) IMMEDIATE SOURCE:

318 (B) CLONE: Preferred non-translated 5' sequence for

## Patent Application US/07/659,408

319 animal cells  
320  
321 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:  
322 AGCTTGCCGC CACT 14  
323  
324 (2) INFORMATION FOR SEQ ID NO:6:  
325  
326 (i) SEQUENCE CHARACTERISTICS:  
327 (A) LENGTH: 906 base pairs  
328 (B) TYPE: nucleic acid  
329 (C) STRANDEDNESS: single  
330 (D) TOPOLOGY: linear  
331  
332 (ii) MOLECULE TYPE: DNA (genomic)  
333  
334 (iii) HYPOTHETICAL: NO  
335  
336  
337 (vii) IMMEDIATE SOURCE:  
338 (B) CLONE: Preferred sequence for expression in animal  
339 cells  
340  
341  
342 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:  
343  
344 ATGTCCGCAG TAAAAGCAGC CCGCTACGGC AAGGACAATG TCCGCGTCTA CAAGGTTTAC 60  
345  
346 AAGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTGT GCTTCTGGAG 120  
347  
348 GGTGAGATTG AGACCTCTTA CACCAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC 180  
349  
350 ATTAAGAACA CCATTTACAT CACCGCCAAG CAGAACCCCG TTAATCCTCC CGAGCTGTTC 240  
351  
352 GGCTCCATCC TGGGCACACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC 300  
353  
354 AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC 360  
355  
356 TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCTGA GGGCAAGGGC 420  
357  
358 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC 480  
359  
360 TGGGGCTTCC TGCCTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC 540  
361  
362 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAAGT GACTCCAGGA GGTCCGCTCG 600  
363  
364 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT 660  
365  
366 GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG 720  
367  
368 GCGCGCCAGC AGCTGATCGA GACTGTCTGAG TACTCGTTGC CTAACAAGCA CTATTTTCGAA 780  
369  
370 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 840  
371

## Patent Application US/07/659,408

372 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 900

373

374 AAATTG

906

375

376 (2) INFORMATION FOR SEQ ID NO:7:

377

(i) SEQUENCE CHARACTERISTICS:

379 (A) LENGTH: 23 base pairs

380 (B) TYPE: nucleic acid

381 (C) STRANDEDNESS: single

382 (D) TOPOLOGY: linear

383

384 (ii) MOLECULE TYPE: DNA (genomic)

385 (iii) HYPOTHETICAL: NO

386

387

388 (vii) IMMEDIATE SOURCE:

389 (B) CLONE: reverse transcription primer

390

391

392 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

393

394 GATCCGGGCC CTTTTTTTTT TTT

23

395

396 (2) INFORMATION FOR SEQ ID NO:8:

397

(i) SEQUENCE CHARACTERISTICS:

399 (A) LENGTH: 10 amino acids

400 (B) TYPE: amino acid

401 (C) STRANDEDNESS: single

402 (D) TOPOLOGY: linear

403

404 (ii) MOLECULE TYPE: peptide

405

406 (iii) HYPOTHETICAL: NO

407

408

409 (vii) IMMEDIATE SOURCE:

410 (B) CLONE: Hydrolysis product T 17

411

412

413 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

414

415 Asn Val Gln Val Asp Val Val Glu Gly Lys

416 1 5 10

417

418 (2) INFORMATION FOR SEQ ID NO:9:

419

(i) SEQUENCE CHARACTERISTICS:

421 (A) LENGTH: 8 amino acids

422 (B) TYPE: amino acid

423 (C) STRANDEDNESS: single

424 (D) TOPOLOGY: linear



## Patent Application US/07/659,408

425  
426 (ii) MOLECULE TYPE: peptide  
427  
428 (iii) HYPOTHETICAL: NO  
429  
430 (vii) IMMEDIATE SOURCE:  
431 (B) CLONE: Hydrolysis product T 20  
432

433  
434 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:  
435  
436 Asn Phe Ser Gly Leu Gln Glu Val  
437 1 5  
438

439 (2) INFORMATION FOR SEQ ID NO:10:  
440

441 (i) SEQUENCE CHARACTERISTICS:  
442 (A) LENGTH: 6 amino acids  
443 (B) TYPE: amino acid  
444 (C) STRANDEDNESS: single  
445 (D) TOPOLOGY: linear  
446

447 (ii) MOLECULE TYPE: peptide  
448

449 (iii) HYPOTHETICAL: NO  
450

451  
452 (vii) IMMEDIATE SOURCE:  
453 (B) CLONE: Hydrolysis product T 23  
454

455  
456 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:  
457

458 Phe Asp Ala Thr Trp Ala  
459 1 5  
460

461 (2) INFORMATION FOR SEQ ID NO:11:  
462

463 (i) SEQUENCE CHARACTERISTICS:  
464 (A) LENGTH: 8 amino acids  
465 (B) TYPE: amino acid  
466 (C) STRANDEDNESS: single  
467 (D) TOPOLOGY: linear  
468

469 (ii) MOLECULE TYPE: peptide  
470

471 (iii) HYPOTHETICAL: NO  
472

473  
474 (vii) IMMEDIATE SOURCE:  
475 (B) CLONE: Hydrolysis product T 27  
476  
477

## Patent Application US/07/659,408

478 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:  
479  
480 His Tyr Phe Glu Ile Asp Leu Ser  
481 1 5  
482

483 (2) INFORMATION FOR SEQ ID NO:12:  
484  
485 (i) SEQUENCE CHARACTERISTICS:  
486 (A) LENGTH: 13 amino acids  
487 (B) TYPE: amino acid  
488 (C) STRANDEDNESS: single  
489 (D) TOPOLOGY: linear  
490  
491 (ii) MOLECULE TYPE: peptide  
492  
493 (iii) HYPOTHETICAL: NO  
494  
495  
496 (vii) IMMEDIATE SOURCE:  
497 (B) CLONE: Hydrolysis product T 28  
498  
499

500 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:  
501  
502 Ile Leu Ser Thr Asp Val Asp Ala Thr Trp Gln Trp Lys  
503 1 5 10  
504

505 (2) INFORMATION FOR SEQ ID NO:13:  
506  
507 (i) SEQUENCE CHARACTERISTICS:  
508 (A) LENGTH: 11 amino acids  
509 (B) TYPE: amino acid  
510 (C) STRANDEDNESS: single  
511 (D) TOPOLOGY: linear  
512  
513 (ii) MOLECULE TYPE: peptide  
514  
515 (iii) HYPOTHETICAL: NO  
516  
517  
518 (vii) IMMEDIATE SOURCE:  
519 (B) CLONE: Hydrolysis product T 29  
520  
521

522 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:  
523  
524 His Tyr Phe Glu Ile Asp Leu Ser Trp His Lys  
525 1 5 10  
526

527 (2) INFORMATION FOR SEQ ID NO:14:  
528  
529 (i) SEQUENCE CHARACTERISTICS:  
530 (A) LENGTH: 11 amino acids

## Patent Application US/07/659,408

531 (B) TYPE: amino acid  
532 (C) STRANDEDNESS: single  
533 (D) TOPOLOGY: linear  
534  
535 (ii) MOLECULE TYPE: peptide  
536  
537 (iii) HYPOTHETICAL: NO  
538  
539  
540 (vii) IMMEDIATE SOURCE:  
541 (B) CLONE: Hydrolysis product T 31  
542  
543  
544 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:  
545  
546 Ser Thr Asn Ser Gln Phe Trp Gly Phe Leu Arg  
547 1 5 10  
548  
549 (2) INFORMATION FOR SEQ ID NO:15:  
550  
551 (i) SEQUENCE CHARACTERISTICS:  
552 (A) LENGTH: 16 amino acids  
553 (B) TYPE: amino acid  
554 (C) STRANDEDNESS: single  
555 (D) TOPOLOGY: linear  
556  
557 (ii) MOLECULE TYPE: peptide  
558  
559 (iii) HYPOTHETICAL: NO  
560  
561  
562 (vii) IMMEDIATE SOURCE:  
563 (B) CLONE: Hydrolysis product T 32  
564  
565  
566 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:  
567  
568 Gln Asn Pro Val Thr Pro Pro Glu Leu Phe Gly Ser Ile Leu Gly Thr  
569 1 5 10 15  
570  
571  
572 (2) INFORMATION FOR SEQ ID NO:16:  
573  
574 (i) SEQUENCE CHARACTERISTICS:  
575 (A) LENGTH: 16 amino acids  
576 (B) TYPE: amino acid  
577 (C) STRANDEDNESS: single  
578 (D) TOPOLOGY: linear  
579  
580 (ii) MOLECULE TYPE: peptide  
581  
582 (iii) HYPOTHETICAL: NO  
583

## Patent Application US/07/659,408

584

585 (vii) IMMEDIATE SOURCE:

586 (B) CLONE: Hydrolysis product T 33

587

588

589 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

590

591 Gln Asn Pro Val Thr Pro Pro Glu Leu Phe Gly Ser Ile Leu Gly Thr

592 1 5 10 15

593

594

595 (2) INFORMATION FOR SEQ ID NO:17:

596

597 (i) SEQUENCE CHARACTERISTICS:

598 (A) LENGTH: 25 amino acids

599 (B) TYPE: amino acid

600 (C) STRANDEDNESS: single

601 (D) TOPOLOGY: linear

602

603 (ii) MOLECULE TYPE: peptide

604

605 (iii) HYPOTHETICAL: NO

606

607

608 (vii) IMMEDIATE SOURCE:

609 (B) CLONE: Hydrolysis product V 1

610

611

612 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

613

614 Tyr Ser Leu Pro Asn Lys Ala Asp Asn Ser Val Ile Val Asp Thr Asp

615 1 5 10 15

616

617 Ser Ile Lys Asn Thr Ile Tyr Ile Thr

618 20 25

619

620 (2) INFORMATION FOR SEQ ID NO:18:

621

622 (i) SEQUENCE CHARACTERISTICS:

623 (A) LENGTH: 16 amino acids

624 (B) TYPE: amino acid

625 (C) STRANDEDNESS: single

626 (D) TOPOLOGY: linear

627

628 (ii) MOLECULE TYPE: peptide

629

630 (iii) HYPOTHETICAL: NO

631

632

633 (vii) IMMEDIATE SOURCE:

634 (B) CLONE: Hydrolysis product V 2

635

636

## Patent Application US/07/659,408

637 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

638  
639 Val Thr Leu Lys Thr Phe Ala Glu Asp Asn Ser Ala Ser Val Gln Ala  
640 1 5 10 15  
641

642 (2) INFORMATION FOR SEQ ID NO:19:

643  
644 (i) SEQUENCE CHARACTERISTICS:  
645 (A) LENGTH: 24 amino acids  
646 (B) TYPE: amino acid  
647 (C) STRANDEDNESS: single  
648 (D) TOPOLOGY: linear  
649

650 (ii) MOLECULE TYPE: peptide

651  
652 (iii) HYPOTHETICAL: NO  
653

654  
655 (vii) IMMEDIATE SOURCE:  
656 (B) CLONE: Hydrolysis product V 3  
657  
658

659 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

660  
661 Thr Ser Tyr Thr Lys Ala Asp Asn Ser Val Ile Val Asp Thr Asp Ser  
662 1 5 10 15  
663  
664 Ile Lys Asn Thr Ile Tyr Ile Thr  
665 20  
666

667 (2) INFORMATION FOR SEQ ID NO:20:

668  
669 (i) SEQUENCE CHARACTERISTICS:  
670 (A) LENGTH: 28 amino acids  
671 (B) TYPE: amino acid  
672 (C) STRANDEDNESS: single  
673 (D) TOPOLOGY: linear  
674

675 (ii) MOLECULE TYPE: peptide

676  
677 (iii) HYPOTHETICAL: NO  
678

679  
680 (vii) IMMEDIATE SOURCE:  
681 (B) CLONE: Hydrolysis product V 5  
682  
683

684 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

685  
686 Gly Lys Gly Ile Asp Ile Lys Ser Ser Leu Ser Gly Leu Thr Val Leu  
687 1 5 10 15  
688  
689 Lys Ser Thr Asn Ser Gln Phe Trp Gly Phe Leu Arg

#6

## Patent Application US/07/659,408

690 20 25

691

692 (2) INFORMATION FOR SEQ ID NO:21:

693

694 (i) SEQUENCE CHARACTERISTICS:

695 (A) LENGTH: 17 amino acids

696 (B) TYPE: amino acid

697 (C) STRANDEDNESS: single

698 (D) TOPOLOGY: linear

699

700 (ii) MOLECULE TYPE: peptide

701

702 (iii) HYPOTHETICAL: NO

703

704

705 (vii) IMMEDIATE SOURCE:

706 (B) CLONE: Hydolysis product V 6

707

708

709 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

710

711 Gly Lys Gly Ile Asp Ile Lys Ser Ser Leu Ser Gly Leu Thr Val Leu

712 1 5 10 15

713

714 Lys

715

716

717 (2) INFORMATION FOR SEQ ID NO:22:

718

719 (i) SEQUENCE CHARACTERISTICS:

720 (A) LENGTH: 1236 base pairs

721 (B) TYPE: nucleic acid

722 (C) STRANDEDNESS: single

723 (D) TOPOLOGY: linear

724

725 (ii) MOLECULE TYPE: DNA (genomic)

726

727 (iii) HYPOTHETICAL: NO

728

729

730 (vii) IMMEDIATE SOURCE:

731 (B) CLONE: Fragment 3

732

733

734 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

735

736 GATCCGCGGA AGCATAAAGT GTAAAGCCTG GGGTGCCTAA TGAGTGAGCT AACTTACATT 60

737

738 AATTGCGTTG CGCTCACTGC CCGCTTTCCA GTCGGGAAAC CTGTCGTGCC AGCTGCATTA 120

739

740 ATGAATCGGC CAACGCGCGG GGAGAGGCGG TTTGCGTATT GGGCGCCAGG GTGGTTTTTC 180

741

742 TTTTCACCAG TGAGACGGGC AACAGCTGAT TGCCCTTCAC CGCCTGGCCC TGAGAGAGTT 240

## Patent Application US/07/659,408

743  
744 GCAGCAAGCG GTCCACGCTG GTTTGCCCCA CCACCCGAAA ATCCTGTTTG ATGGTGGTTA 300  
745  
746 ACGGCGGGAT ATAACATGAG CTGTCTTCGG TATCGTCGTA TCCCACTACC GAGATATCCG 360  
747  
748 CACCAACGCG CAGCCCGGAC TCGGTAATGG CGCGCATTCG GCCCAGCGCC ATCTGATCGT 420  
749  
750 TGGCAACCAG CATCGCAGTG GGAACGATGC CCTCATTCAG CATTTGCATG GTTTGTTGAA 480  
751  
752 AACC GGACAT GGC ACTCCAG TCGCCTTCCC GTTCCGCTAT CGGCTGAATT TGATTGCGAG 540  
753  
754 TGAGATATTT ATGCCAGCCA GCCAGACGCA GACGCGCCGA GACAGAACTT AATGGGCCCCG 600  
755  
756 CTAACAGCGC GATTTGCTGG TGACCCAATG CGACCAGATG CTCCACGCCC AGTCGCGTAC 660  
757  
758 CGTCTTCATG GGAGAAAATA ATACTGTTGA TGGGTGTCTG GTCAGAGACA TCAAGAAATA 720  
759  
760 ACGCCGGAAC ATTAGTGAGC GCAGCTTCCA CAGCAATGGC ATCCTGGTCA TCCAGCGGAT 780  
761  
762 AGTTAATGAT CAGCCCACTG ACGCGTTGCG CGAGAAGATT GTGCACCGCC GCTTTACAGG 840  
763  
764 CTTGACGCGC GCTTCGTTCT ACCATCGACA CCACCACGCT GGCACCCAGT TGATCGGCGC 900  
765  
766 GAGATTTAAT CGCCGCGACA ATTTGCGACG GCGCGTGCAG GGCCAGACTG GAGGTGGCAA 960  
767  
768 CGCCAATCAG CAACGACTGT TTGCCCAGCA GTTGTGTGTC CACGCGGTTG GGAATGTAAT 1020  
769 TCAGCTCCGC CATCGCCGCT TCCACTTTTT CCCGCGTTTT CGCAGAAACG TGGCTGGCCT 1080  
770  
771 GGTTCACCAC GCGGGAAACG GTCTGATAAC AGACACCGGC ATACTCTGCG ACATCGTATA 1140  
772  
773 ACGTTACTGG TTTCACATTC ACCACCCTGA ATTGACTCTC TTCCGGGCGC TATCATGCCA 1200  
774  
775 TACCGCGAAA GGTTTTGCGC CATTCGATGG TGTCCG 1236  
776

## (2) INFORMATION FOR SEQ ID NO:23:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 326 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: NO

## (vii) IMMEDIATE SOURCE:

(B) CLONE: Fragment 4

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

## Patent Application US/07/659,408

796 TCGAGCTGAC TGACCTGTTG CTTATATTAC ATCGATAGCT TAGCGTATAA TGTGTGGAAT 60  
797  
798 TGTGAGCGAT AACAAATTCA CACAGTTTAA CTTTAAGAAG GAGATATACA TATGGCTACC 120  
799  
800 GGATCCCGGA CTAGTCTGCT CCTGGCTTTT GGCCTGCTCT GCCTGCCCTG GCTTCAAGAG 180  
801  
802 GGCAGTGCCT TCCCAACCAT TCCCTTATCT AGACTTTTGG ACAACGCTAT GCTCCGCGCC 240  
803  
804 CATCGTCTGC ACCAGCTGGC CTTTGACACC TACCAGGAGT TTGAAGAAGC CTATATCCCA 300  
805 AAGGAACAGA AGTATTCATT CCTGCA 326  
806

## 807 (2) INFORMATION FOR SEQ ID NO:24:

808

## 809 (i) SEQUENCE CHARACTERISTICS:

810 (A) LENGTH: 74 base pairs

811 (B) TYPE: nucleic acid

812 (C) STRANDEDNESS: single

813 (D) TOPOLOGY: linear

814

815 (ii) MOLECULE TYPE: DNA (genomic)

816

817 (iii) HYPOTHETICAL: NO

818

819

820 (vii) IMMEDIATE SOURCE:

821 (B) CLONE: ClaI-NdeI fragment

822

823

824 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

825

826 CGATAGCGTA TAATGTGTGG AATTGTGAGC GGATAACAAT TTCACACAGT TTTTCGCGAA 60

827

828 GAAGGAGATA TACA 74

829

## 830 (2) INFORMATION FOR SEQ ID NO:25:

831

## 832 (i) SEQUENCE CHARACTERISTICS:

833 (A) LENGTH: 163 base pairs

834 (B) TYPE: nucleic acid

835 (C) STRANDEDNESS: single

836 (D) TOPOLOGY: linear

837

838 (ii) MOLECULE TYPE: DNA (genomic)

839

840 (iii) HYPOTHETICAL: NO

841

842

843 (vii) IMMEDIATE SOURCE:

844 (B) CLONE: Synthetic hGH gene end fragment

845

846

847 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

848



## Patent Application US/07/659,408

849 GATCTTCAAG CAGACCTACA GCGACGCACT ACTCAAGAAC TACGGGCTGC TCTACTGCTT 60  
850  
851 CAGGAAGGAC ATGGACAAGG TCGAGACATT CCTGCGCATC GTGCAGTGCC GCTCTGTGGA 120  
852  
853 GGGCAGCTGT GGCTTCTAGT AAGGTACCCT GCCCTACGTA CCA 163  
854  
855 (2) INFORMATION FOR SEQ ID NO:26:  
856  
857 (i) SEQUENCE CHARACTERISTICS:  
858 (A) LENGTH: 48 base pairs  
859 (B) TYPE: nucleic acid  
860 (C) STRANDEDNESS: single  
861 (D) TOPOLOGY: linear  
862  
863 (ii) MOLECULE TYPE: DNA (genomic)  
864  
865 (iii) HYPOTHETICAL: NO  
866  
867  
868 (vii) IMMEDIATE SOURCE:  
869 (B) CLONE: NdeI-AccII synthetic fragment  
870  
871  
872 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:  
873  
874 TATGTCTGCG GTAAAAGCAG CGCGCTACGG CAAGGACAAT GTTCGCGT 48  
875  
876 (2) INFORMATION FOR SEQ ID NO:27:  
877  
878 (i) SEQUENCE CHARACTERISTICS:  
879 (A) LENGTH: 361 base pairs  
880 (B) TYPE: nucleic acid  
881 (C) STRANDEDNESS: single  
882 (D) TOPOLOGY: linear  
883  
884 (ii) MOLECULE TYPE: DNA (genomic)  
885  
886 (iii) HYPOTHETICAL: NO  
887  
888  
889 (vii) IMMEDIATE SOURCE:  
890 (B) CLONE: Plasmid pEMR469 fragment with ADH2 promoter  
891  
892  
893 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:  
894  
895 GGGACGCGTC TCCTCTGCCG GAACACCGGG CATCTCCAAC TTATAAGTTG GAGAAATAAG 60  
896  
897 AGAATTTTCAG ATTGAGAGAA TGAAAAAAAA AAAAAAAAAA AAGGCAGAGG AGAGCATAGA 120  
898  
899 AATGGGGTTC ACTTTTTGGT AAAGCTATAG CATGCCTATC ACATATAAAT AGAGTGCCAG 180  
900  
901 TAGCGACTTT TTTCACACTC GAGATACTCT TACTACTGCT CTCTTGTTGT TTTTATCACT 240

## Patent Application US/07/659,408

902  
903 TCTTGTCTTCT TCTTTGGTAA ATAGAATATC AAGCTACAAA AAGCATACAA TCAACTATCA 300  
904  
905 ACTATTAACT ATATCGATAC CATATGGATC CGTCGACTCT AGAGGATCGT CGACTCTAGA 360  
906  
907 G 361  
908  
909 (2) INFORMATION FOR SEQ ID NO:28:  
910  
911 (i) SEQUENCE CHARACTERISTICS:  
912 (A) LENGTH: 58 base pairs  
913 (B) TYPE: nucleic acid  
914 (C) STRANDEDNESS: single  
915 (D) TOPOLOGY: linear  
916  
917 (ii) MOLECULE TYPE: DNA (genomic)  
918  
919 (iii) HYPOTHETICAL: NO  
920  
921  
922 (vii) IMMEDIATE SOURCE:  
923 (B) CLONE: Fragment C  
924  
925  
926 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:  
927  
928 CGATATACAC AATGTCTGCT GTTAAGGCTG CTAGATACGG TAAGGACAAC GTTAGAGT 58  
929  
930 (2) INFORMATION FOR SEQ ID NO:29:  
931  
932 (i) SEQUENCE CHARACTERISTICS:  
933 (A) LENGTH: 1013 base pairs  
934 (B) TYPE: nucleic acid  
935 (C) STRANDEDNESS: single  
936 (D) TOPOLOGY: linear  
937  
938 (ii) MOLECULE TYPE: DNA (genomic)  
939  
940 (iii) HYPOTHETICAL: NO  
941  
942  
943 (vii) IMMEDIATE SOURCE:  
944 (B) CLONE: Fragment D  
945  
946  
947 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:  
948  
949 CTACAAGGTT CACAAGGACC ACAAGACCGG TGTCCAGACG GTGTACGAGA TGACCGTCTG 60  
950  
951 TGTGCTTCTG GAGGGTGAGA TTGAGACCTC TTACACCAAG GCCGACAACA GCGTCATTGT 120  
952  
953 CGCAACCGAC TCCATTAAAG ACACCATTTA CATCACCGCC AAGCAGAACC CCGTTACTCC 180  
954

## Patent Application US/07/659,408

955 TCCCGAGCTG TTCGGCTCCA TCCTGGGCAC ACACCTTCAT TGAGAAGTAC AACACATCCA 240  
956  
957 TGCCGCTCAC GTCAACATTG TCTGCCACCG CTGGACCCGG ATGGACATTG ACGGCAAGCC 300  
958  
959 ACACCCTCAC TCCTTCATCC GCGACAGCGA GGAGAAGCGG AATGTGCAGG TGGACGTGGT 360  
960  
961 CGAGGGCAAG GGCATCGATA TCAAGTCGTC TCTGTCCGGC CTGACCGTGC TGAAGAGCAC 420  
962  
963 CAACTCGCAG TTCTGGGGCT TCCTGCGTGA CGAGTACACC ACACCTAAGG AGACCTGGGA 480  
964  
965 CCGTATCCTG AGCACCGACG TCGATGCCAC TTGGCAGTGG AAGAATTTCA GTGGACTCCA 540  
966  
967 GGAGGTCCGC TCGCACGTGC CTAAGTTCGA TGCTACCTGG GCCACTGCTC GCGAGGTCAC 600  
968  
969 TCTGCCGACT TTTGCTGAAG ATAACAGTGC CAGCGTGCAG GCCACTATGT ACAAGATGGC 660  
970  
971 AGAGCAAATC CTGGCGCGCC AGCAGCTGAT CGAGACTGTC GAGTACTCGT TGCCTAACAA 720  
972  
973 GCACTATTTT GAAATCGACC TGAGCTGGCA CAAGGGCCTC CAAAACACCG GCAAGAACGC 780  
974  
975 CGAGGTCTTC GCTCCTCAGT CGGACCCCAA CGGTCTGATC AAGTGTACCG TCGGCCGGTC 840  
976  
977 CTCTCTGAAG TCTAAATTGT AAACCAACAT GATTCTCACG TTCCGGAGTT TCCAAGGCAA 900  
978  
979 ACTGTATATA GTCTGGGATA GGGTATAGCA TTCATTCACT TGTTTTTTTAC TTCCAAAAAA 960  
980  
981 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAGGGC CCG 1013  
982  
983 (2) INFORMATION FOR SEQ ID NO:30:  
984  
985 (i) SEQUENCE CHARACTERISTICS:  
986 (A) LENGTH: 207 base pairs  
987 (B) TYPE: nucleic acid  
988 (C) STRANDEDNESS: single  
989 (D) TOPOLOGY: linear  
990  
991 (ii) MOLECULE TYPE: DNA (genomic)  
992  
993 (iii) HYPOTHETICAL: NO  
994  
995  
996 (vii) IMMEDIATE SOURCE:  
997 (B) CLONE: Synthetic GAL7 fragment  
998  
999  
1000 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:  
1001  
1002 CGCGTCTATA CTTCGGAGCA CTGTTGAGCG AAGGCTCATT AGATATATTT TCTGTCATTT 60  
1003  
1004 TCCTTAACCC AAAAATAAGG GAGAGGGTCC AAAAAGCGCT CGGACAACCTG TTGACCGTGA 120  
1005  
1006 TCCGAAGGAC TGGCTATACA GTGTTCAACA AATAGCCAAG CTGAAAATAA TGTGTAGCCT 180  
1007

## Patent Application US/07/659,408

1008 TTAGCTATGT TCAGTTAGTT TGGCATG 207  
1009  
1010 (2) INFORMATION FOR SEQ ID NO:31:  
1011  
1012 (i) SEQUENCE CHARACTERISTICS:  
1013 (A) LENGTH: 23 base pairs  
1014 (B) TYPE: nucleic acid  
1015 (C) STRANDEDNESS: single  
1016 (D) TOPOLOGY: linear  
1017  
1018 (ii) MOLECULE TYPE: DNA (genomic)  
1019  
1020 (iii) HYPOTHETICAL: NO  
1021  
1022 (vii) IMMEDIATE SOURCE:  
1023 (B) CLONE: Modified XbaI-MluI adapter  
1024  
1025 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:  
1026  
1027 CTAGGCTAGC GGGCCCCGCAT GCA 23  
1028  
1029 (2) INFORMATION FOR SEQ ID NO:32:  
1030  
1031 (i) SEQUENCE CHARACTERISTICS:  
1032 (A) LENGTH: 422 base pairs  
1033 (B) TYPE: nucleic acid  
1034 (C) STRANDEDNESS: single  
1035 (D) TOPOLOGY: linear  
1036  
1037 (ii) MOLECULE TYPE: DNA (genomic)  
1038  
1039 (iii) HYPOTHETICAL: NO  
1040  
1041  
1042 (vii) IMMEDIATE SOURCE:  
1043 (B) CLONE: Plasmid pSE1 "site binding to HindIII"  
1044 fragment  
1045  
1046  
1047 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:  
1048  
1049 AGCTGGCTCG CATCTCTCCT TCACGCGCCC GCCGCCCTAC CTGAGGCCGC CATCCACGCC 60  
1050  
1051 GGTGAGTCGC GTTCTGCCGC CTCCCGCCTG TGGTGCTTCC TGAAGTGCCT CCGCCGTCTA 120  
1052  
1053 GGTAGGCTCC AAGGGAGCCG GACAAAGGCC CGGTCTCGAC CTGAGCTCTA AACTTACCTA 180  
1054  
1055 GACTCAGCCG GCTCTCCACG CTTTGCCTGA CCCTGCTTGC TCAACTCTAC GTCTTTGTTT 240  
1056  
1057 CGTTTTCTGT TCTGCGCCGT TACAACTTCA AGGTATGCGC TGGGACCTGG CAGGCGGCAT 300  
1058  
1059 CTGGGACCCC TAGGAAGGGC TTGGGGGTCC TCGTGCCCAA GGCAGGGAAC ATAGTGGTCC 360  
1060

## Patent Application US/07/659,408

1061 CAGGAAGGGG AGCAGAGGCA TCAGGGTGTCT CACTTTGTCT CCGCAGCTCC TGAGCCTGCA 420  
1062  
1063 GA 422  
1064  
1065 (2) INFORMATION FOR SEQ ID NO:33:  
1066  
1067 (i) SEQUENCE CHARACTERISTICS:  
1068 (A) LENGTH: 77 base pairs  
1069 (B) TYPE: nucleic acid  
1070 (C) STRANDEDNESS: single  
1071 (D) TOPOLOGY: linear  
1072  
1073 (ii) MOLECULE TYPE: DNA (genomic)  
1074 (iii) HYPOTHETICAL: NO  
1075  
1076  
1077 (vii) IMMEDIATE SOURCE:  
1078 (B) CLONE: Synthetic HindIII-"site binding to BamHI"  
1079 fragment  
1080  
1081  
1082 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:  
1083  
1084 AGCTTGTCGA CTAATACGAC TCACTATAGG GCGGCCGCGG GCCCTGCAG GAATTCGGAT 60  
1085  
1086 CCCCCGGGTG ACTGACT 77  
1087  
1088 (2) INFORMATION FOR SEQ ID NO:34:  
1089 (i) SEQUENCE CHARACTERISTICS:  
1090 (A) LENGTH: 61 base pairs  
1091 (B) TYPE: nucleic acid  
1092 (C) STRANDEDNESS: single  
1093 (D) TOPOLOGY: linear  
1094  
1095 (ii) MOLECULE TYPE: DNA (genomic)  
1096  
1097 (iii) HYPOTHETICAL: NO  
1098  
1099  
1100 (vii) IMMEDIATE SOURCE:  
1101 (B) CLONE: Synthetic HindIII-AccI fragment  
1102  
1103  
1104 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:  
1105  
1106 AGCTTGCCGC CACTATGTCC GCAGTAAAAG CAGCCCCTA CGGCAAGGAC AATGTCCGCG 60  
1107  
1108 T 61  
1109  
1110 (2) INFORMATION FOR SEQ ID NO:35:  
1111  
1112 (i) SEQUENCE CHARACTERISTICS:  
1113 (A) LENGTH: 920 base pairs

## Patent Application US/07/659,408

1114 (B) TYPE: nucleic acid  
1115 (C) STRANDEDNESS: single  
1116 (D) TOPOLOGY: linear  
1117  
1118 (ii) MOLECULE TYPE: DNA (genomic)  
1119  
1120 (iii) HYPOTHETICAL: NO  
1121  
1122  
1123 (vii) IMMEDIATE SOURCE:  
1124 (B) CLONE: HindIII-SnaBI fragment  
1125  
1126  
1127 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:  
1128  
1129 AGCTTGCCGC CACTATGTCC GCAGTAAAAG CAGCCCGCTA CGGCAAGGAC AATGTCCGCG 60  
1130  
1131 TCTACAAGGT TCACAAGGAC GAGAAGACCG GTGTCCAGAC GGTGTACGAG ATGACCGTCT 120  
1132  
1133 GTGTGCTTCT GGAGGGTGAG ATTGAGACCT CTTACACCAA GGCCGACAAC AGCGTCATTG 180  
1134  
1135 TCGCAACCGA CTCCATTAAAG AACACCATTT ACATCACCGC CAAGCAGAAC CCCGTTACTC 240  
1136  
1137 CTCCCGAGCT GTTCGGCTCC ATCCTGGGCA CACACTTCAT TGAGAAGTAC AACCACATCC 300  
1138  
1139 ATGCCGCTCA CGTCAACATT GTCTGCCACC GCTGGACCCG GATGGACATT GACGGCAAGC 360  
1140  
1141 CACACCCTCA CTCCTTCATC CGCGACAGCG AGGAGAAGCG GAATGTGCAG GTGGACGTGG 420  
1142  
1143 TCGAGGGCAA GGGCATCGAT ATCAAGTCGT CTCTGTCCGG CCTGACCGTG CTGAAGAGCA 480  
1144  
1145 CCAACTCGCA GTTCTGGGGC TTCCTGCGTG ACGAGTACAC CAACTTAAG GAGACCTGGG 540  
1146  
1147 ACCGTATCCT GAGCACCGAC GTCGATGCCA CTTGGCAGTG GAAGAATTTC AGTGGACTCC 600  
1148  
1149 AGGAGGTCCG CTCGCACGTG CCTAAGTTCG ATGCTACCTG GGCCACTGCT CGCGAGGTCA 660  
1150  
1151 CTCTGAAGAC TTTTGCTGAA GATAACAGTG CCAGCGTGCA GGCCACTATG TACAAGATGG 720  
1152  
1153 CAGAGCAAAT CCTGGCGCGC CAGCAGCTGA TCGAGACTGT CGAGTACTCG TTGCCTAACA 780  
1154  
1155 AGCACTATTT CGAAATCGAC CTGAGCTGGC ACAAGGGCCT CAAAACACC GGCAAGAACG 840  
1156  
1157 CCGAGGTCTT CGCTCCTCAG TCGGACCCA ACGGTCTGAT CAAGTGTACC GTCGGCCGGT 900  
1158  
1159 CCTCTCTGAA GTCTAAATTG 920

PAGE: 1

SEQUENCE VERIFICATION REPORT  
PATENT APPLICATION US/07/659,408

DATE: 10/23/91  
TIME: 11:14:19

LINE ERROR

ORIGINAL TEXT

38 Wrong application Serial Number  
219 Response Exceeds Line Limitations  
268 Response Exceeds Line Limitations  
319 Response Exceeds Line Limitations  
339 Response Exceeds Line Limitations  
1044 Response Exceeds Line Limitations  
1079 Response Exceeds Line Limitations

(A) APPLICATION NUMBER: US 07/659,408

prokaryotes  
eukaryotes  
animal cells  
cells  
fragment  
fragment

(OK)

PAGE: 1

SEQUENCE MISSING ITEM REPORT  
PATENT APPLICATION US/07/659,408

DATE: 10/23/91  
TIME: 11:14:19

MANDATORY IDENTIFIER THAT WAS NOT FOUND

PRIOR APPLICATION DATA  
APPLICATION NUMBER  
FILING DATE



PAGE: 1

SEQUENCE CORRECTION REPORT  
PATENT APPLICATION US/07/659,408

DATE: 10/23/91  
TIME: 11:14:19

LINE ORIGINAL TEXT

CORRECTED TEXT